DOCUMENT RESUME

ED 265 801 HE 019 052

AUTHOR Gleason, Maryellen
TITLE Getting a Perspective on Student Evaluation.
INSTITUTION American Association for Higher Education, Washington, D.C.
PUB DATE Feb 86
NOTE 5p.
AVAILABLE FROM American Association for Higher Education, One Dupont Circle, Suite 600, Washington, DC 20036.
PUB TYPE Collected Works - Serials (022) -- Viewpoints (120)
JOURNAL CIT AAHE Bulletin; p10-13 Feb 1986

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; *College Faculty; College Instruction; College Students; Grades (Scholastic); Higher Education; *Student Evaluation of Teacher Performance; *Teacher Characteristics; *Teacher Effectiveness; Teacher Student Relationship

ABSTRACT

Issues concerning student evaluation of instruction are addressed. Answers to the following questions are provided: Are students in a position to identify the factors that make instruction effective? Do they know what makes teaching good? If ratings are high, does that mean students are learning more? Is there a relationship between instructor ratings and student learning? Do student evaluations of instruction agree with judgments made by faculty and others? Do student ratings tend to be lower in required courses where prior interest in the subject matter may be low? Are student ratings lower in larger classes? Do student ratings change over time? Do students change their minds about effective and ineffective instructors after graduation? Are student evaluation results different depending on when they are completed during a course: Is there a "best" time to administer evaluations? Do comparative data that rank instructors against each other or measure them against a set of norms tell faculty what they need to know in order to improve instruction? To what extent are student assessments of teaching effectiveness a function of an instructor's wit and personality? Is there a relationship between grades and ratings? Publications that provide additional information are cited for each area. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

GETTING A PERSPECTIVE ON STUDENT EVALUATION

by Maryellen Gleason

ED265801

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

AAHE

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Maryellen Gleason is head of the Instructional Development Program at The Pennsylvania State University.

HE019052

Student evaluation of instruction is a topic about which strong opinions are voiced—often sounding like facts. In truth, the collected folk wisdom in most institutions as to the "facts" of evaluation is not always as current or correct as it could be. A number of reasons make this so, not the least being the need of faculty to account for their individual evaluation results.

The issue is more complicated than that, however. Research on student evaluation has been extensive, but finding any single answer calls for a trek through a vast array of accumulated literature. Moreover, the trail winds through a labyrinth of complicated research designs and statistical analyses. It sometimes dead ends, requiring retraced steps or frustrated acknowledgement that the answer has not yet been reached.

Such is the circuitous search for truth in almost any area. And although the challenge to reach for answers spurs experienced researchers on, they are not especially concerned that curious practitioners be able to follow the trails they blaze.

All of this explains why, in the given case, the "facts" about student evaluation frequently derive not from data but from experience and/or the accumulated wisdom of like-minded academicians.

That need not be so. Research answers can be paraphrased and reformulated so that faculty and administrators can be better informed as to what student evaluation can and cannot document about instruction. The objective ought not to be to equip educational practitioners to scale the peak but to get them to a vantage point high enough to see and use what is known about this particular phenomenon.

Achieving a vantage point like that is a desirable goal, but where that place is must be clearly understood. The position provides a perspective that is distant; reformulated answers and explanations obscure details. From this perspective, answers must be taken realizing a certain "imprecision" exists, inherently.

In addition, the peak that student evaluation researchers are challenging must be seen for what it is; when students evaluate instructors they are reporting judgments based on their *perceptions* of what is happening in a classroom. It is a messy sort of science. The dynamic interplay of variables that unfold and interact (often idiosyncratically) in any teaching-learning situation is not well understood or reliably predictable. The ability of even the most sophisticated evaluation instruments, even when analyzed by the most complicated statistical tests, is still a crude attempt to document the effects of instruction. Granted, the data provide a glimpse of what goes on, but they do not tell the whole story.

The following questions and answers attempt to bring educational practitioners to a vantage point that affords them this more accurate and complete perspective. Substantiation is offered for each answer. Sources cited are but a fraction of the number available; those selected represent some of the most comprehensive, current, and/or most widely referenced primary sources relevant to the question.

ARE STUDENTS IN A POSITION TO IDENTIFY THE FACTORS THAT MAKE INSTRUCTION EFFECTIVE? DO THEY KNOW WHAT MAKES TEACHING GOOD?

The question is difficult to answer because no universally accepted or agreed upon standard, criterion, or description of effective teaching exists. However, for several decades now, researchers have been asking students, faculty, administrators, and alumni to identify characteristics of effective instruction and in spite of diverse research approaches, certain characteristics keep appearing on the lists. Moreover, in general, students agree with each other as to the factors that contribute positively to instructional impact. So—the answer is "yes" with the caveat that what constitutes effective instruction has yet to be definitively determined.

- For a well organized summary and review of 49 studies on college student's views of teaching effectiveness see: Feldman, K. A. "The Superior College Teacher From the Students' View." *Research in Higher Education*, 5 (1976), 243-88.
- For a sample of study results identifying characteristics of effective instruction, see a table that succinctly summarizes research across a number of decades and constituencies in: Miller, R. *Developing Programs for Faculty Evaluation*. San Francisco: Jossey-Bass, 1975, pp. 32-3.

IF RATINGS ARE HIGH, DOES THAT MEAN STUDENTS ARE LEARNING MORE? IS THERE A RELATIONSHIP BETWEEN INSTRUCTOR RATINGS AND STUDENT LEARNING?

Yes—not an unqualified yes, as research findings are mixed, but taken as a whole, they confirm that teachers rated as effective by students are generally those teachers whose students achieve more, meaning their final exam scores are higher, they can better apply course materials, and they are more inclined to pursue the subject subsequently. Or, as Cohen concludes, "Students do a pretty good job of distinguishing among teachers on the basis of how much they have learned." (p. 305)

- Centra, J. A. "Student Ratings of Instruction and Their Relationship to Student Learning." *American Educational Research Journal*, 14 (Winter 1977), 17-24.
- A specialized research methodology is used to synthesize 41 separate studies on the relationship between ratings and achievement in: Cohen, P. A. "Student Ratings of Instruction and Student Achievement: A Meta-Analysis of Multisection Validity Studies." *Review of Educational Research*, 51 (Fall 1981), 281-309.
- Marsh, H. W. and Overall, J. U. "Validity of Students' Evaluations of Teaching Effectiveness: Cognitive and Affective Criteria." *Journal of Educational Psychology*, 72 (1980) 468-75.

DO STUDENT EVALUATIONS OF INSTRUCTION AGREE WITH THE JUDGMENTS MADE BY FACULTY AND OTHERS?

Yes, to a substantial degree. Faculty and students agree on the evaluative dimensions—meaning they think the same factors contribute to instructional effectiveness. Faculty and student judgments of the same instructor tend to be in general agreement, as are faculty and student evaluations of a given course.

- Blackburn, R. T. and Clark, M. J. "An Assessment of Faculty Performance: Some Correlates Between Administrator, Colleague, Student and Self-Ratings." *Sociology of Education*, 48 (Spring 1975), 242-56.
- Doyle, K. O. and Crichton, L. I. "Student, Peer and Self Evaluations of College Instructors." *Journal of Educational Psychology*, 70 (1978), 815-26.
- Marsh, H. W.; Overall, J. U.; and Kesler, S. P. "Validity of Student Evaluations of Instructional Effectiveness: A Comparison of Faculty Self-Evaluations and Evaluations by Their Students." *Journal of Educational Psychology*, 71 (1979), 149-60.

DO STUDENT RATINGS TEND TO BE LOWER IN REQUIRED COURSES WHERE PRIOR INTEREST IN THE SUBJECT MATTER MAY BE LOW?

Maybe a little. Research on student evaluations has repeatedly tried to assess the effect of various course and student characteristics on evaluation results to see which (if any) influence the outcome. For example, Marsh looked at 16 variables including workload/difficulty of the course, teacher rank, expected grade, reason for taking the course, class level, among others and found that all 16 variables only accounted for 12 to 14 percent of the variance in student ratings. Among that group of background characteristics, prior interest in the subject was the most influential in terms of its effect on ratings. So, required courses tend to be rated a little (as in very little) lower than electives.

- Marsh, H. W. "The Influence of Student, Course, Instructor Characteristics in Evalua-

tions of University Teaching." *American Educational Research Journal*, 17 (Summer 1980), 219-37.

- Marsh, H. W. and Cooper, T. L. "Prior Subject Interest, Students' Evaluations and Instructional Effectiveness." *Multivariate Behavioral Research*, 16 (January 1981), 83-104.

ARE STUDENT RATINGS LOWER IN LARGER CLASSES?

Maybe a little. Again the answer is not definite. Feldman reviews 30 studies that address the question. About one-third of those report no relationship between class size and ratings. The other two-thirds suggest a modest relationship, i.e. the smaller the class, the higher the ratings. The key is the modesty of the relationship which, depending on the study, accounts for anywhere from one to eight percent of the variance in ratings—definitely not enough to use large courses as an excuse for low ratings.

- For an excellent, succinct summary of an extensive research project on student and course characteristics that includes a numerical breakdown of the effect by class size see: Centra, J. *Determining Faculty Effectiveness*. San Francisco: Jossey-Bass, 1979, pp. 29-30.
- Feldman, K. A. "Course Characteristics and College Students' Ratings of Their Teachers: What We Know and What We Don't." *Research in Higher Education*, 9 (1978), 199-242.

DO STUDENT RATINGS CHANGE OVER TIME? DO STUDENTS CHANGE THEIR MINDS ABOUT EFFECTIVE AND INEFFECTIVE INSTRUCTORS AFTER GRADUATION?

Not usually. Student evaluations change very little even if they are collected some years after students have graduated. It may be that faculty find the opinions of former students more credible, but they are not likely to find those opinions at odds with what they hear from students now.

- Centra, J. A. "The Relationship Between Student and Alumni Ratings of Teachers." *Educational and Psychological Measurement*, 34 (1974), 321-26.
- Marsh, H. W. "The Validity of Students' Evaluations: Classroom Evaluation of Instructors Independently Nominated as Best and Worst Teachers by Graduating Seniors." *American Educational Research Journal*, 14 (Fall 1977), 441-47.

ARE STUDENT EVALUATION RESULTS DIFFERENT DEPENDING ON WHEN THEY ARE COMPLETED DURING A COURSE? IS THERE A "BEST" TIME TO ADMINISTER EVALUATIONS?

It depends—if the goal is to use information from students to determine how a particular course is progressing and what (if any) alterations might be indicated, obviously a time at mid-course can better accomplish those aims than one near the course's conclusion.

- Overall, J. U. and Marsh, H. W. "Mid-Term Feedback From Students: Its Relationship to Instructional Improvement and Students' Cognitive and Affective Outcome." *Journal of Educational Psychology*, 71 (1979), 856-65.

However, if the goal is to produce the kind of data needed to establish one's instructional prowess compared to that of other instructors, research indicates that *when* results are collected appears to make little difference. Students make decisions about an instructor's competence early in the course and remain pretty much committed to those conclusions throughout the term or semester.

- For a table summarizing research findings that documents the extent to which ratings do not change or remain stable across a course, see: Doyle, K. O. *Student Evaluation of Instruction*. Lexington, Massachusetts: Lexington Books, 1975, p. 39.
- Kohlan, R. "A Comparison of Faculty Evaluations Early and Late in the Course." *Journal of Higher Education*, 44 (November 1973), 587-95.
- Frey, P. W. "Validity of Student Instructional Ratings: Does Timing Matter?" *Journal of Higher Education*, 47 (May/June 1976), 327-36.

DO COMPARATIVE DATA THAT RANK INSTRUCTORS AGAINST EACH OTHER OR MEASURE THEM AGAINST A SET OF NORMS TELL FACULTY WHAT THEY NEED TO KNOW IN ORDER TO IMPROVE INSTRUCTION?

No—that kind of information is not particularly effective in modifying instructional practice. Low scores may motivate improvement but that information alone is typically not specific or descriptive enough to enable faculty to know what aspects of the instruction need to be changed. Research indicates that information conducive to improvement needs to be descriptive, not judgmental, needs to occur mid-course, and needs to include some opportunity for interaction about the results.

- Cohen, P. A. "Effectiveness of Student-Rating Feedback for Improving College Instruction: A Meta-Analysis of Findings." *Research in Higher Education*, 13 (1980), 321-41.
- McKeachie, W. J. and Lin, Y. G. *Study 1: Using Student Ratings to Improve Teaching*, Ann Arbor: University of Michigan, 1975, (ERIC document reproduction service #104284).
- Overall, J. U. and Marsh, H. W. "Mid-Term Feedback From Students: Its Relationship to Instructional Improvement in Students' Cognitive and Affective Outcomes." *Journal of Educational Psychology*, 71 (1979), 856-65.
- Rotem, A. and Glasman, N. S. "On Effectiveness of Students' Evaluative Feedback to University Instructors." *Review of Educational Research*, 49 (Summer 1979), 497-511.

TO WHAT EXTENT ARE STUDENT ASSESSMENTS OF TEACHING EFFECTIVENESS A FUNCTION OF AN INSTRUCTOR'S WIT AND PERSONALITY? IS "POPULARITY" THE KEY TO HIGH STUDENT EVALUATIONS? ARE THE "ENTERTAINERS" THE REAL WINNERS IN THE RATINGS GAME?

A rather impressive series of studies in the early 70s did in fact document that student assessments of teaching effectiveness were more a function of the instructor's wit and personality than of the content being presented. However, subsequent researchers (including some in the original group) have been unable to replicate these "Doctor Fox" (the pseudonym for an instructor with wit and personality but no content) studies. The current conclusion, based on instructional settings more closely resembling classrooms, supports the contention that instructor "expressiveness" matters—as it effects ratings—but that effect does not operate independent of content considerations. In other words, students can and do tell the difference.

- Abrami, P. C.; Leventhal, L.; and Perry, R. "Educational Seduction." *Review of Educational Research*, 52 (Fall 1982), 446-64.
- Marsh, H. and Ware, J. "Effects of Expressiveness, Content Coverage and Incentive on Multidimensional Student Rating Scales." *Journal of Educational Psychology*, 74 (1982), 126-34.
- Perry, R. P.; Abrami, P. C.; and Leventhal, L. "Educational Seduction: The Effect of Instructor Expressiveness and Lecture Content on Student Ratings and Achievement." *Journal of Educational Psychology*, 71 (1979), 107-16.

IS THERE A RELATIONSHIP BETWEEN GRADES AND RATINGS? IF GOOD GRADES ARE GIVEN, DO RATINGS RISE PROPORTIONATELY?

Evidence exists on both sides of these questions. However, as the Costin, Greenough, and Menges research review points out, the evidence that says yes is based on correlations that are low. This weakens the claim, especially in light of more recent studies (like the Palmer, Carliner, and Romer study below) that do a better job of disentangling learning, leniency, and evaluation.

- For a summary of early research on both sides of the issue, see: Costin, F.; Greenough, W.; and Menges, R. "Student Ratings of College Teaching: Reliability, Validity, and Usefulness." *Review of Educational Research*, 6 (December 1971), 518-19.
- Palmer, J.; Carliner, G.; and Romer, T. "Leniency, Learning and Evaluations." *Journal of Educational Psychology*, 70 (1978) 855-63.

WHAT ELSE MIGHT BE READ ABOUT THE RESEARCH ON STUDENT EVALUATION?

Two research summaries provide information that elaborates further what is presented here. Both contain excellent bibliographies.

- McKeachie, W. J. "Student Ratings of Faculty: A Reprise." *Academe*, 65 (October 1979), 384-97.
- Overall, J. U. and Marsh, H. W. "Students' Evaluations of Teaching: An Update." *AAHE Bulletin*, Vol. 35, No. 4 (December 1982), 9-13.